NAT'L INST. OF STAND & TECH R.I.C.
A11105 156589

NIST
PUBLICATIONS

NISTIR 6028

Molecular Dynamics and Hydrogen Bonds in Water

Raymond D. Mountain

U.S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards
and Technology
Physical and Chemical Properties Division
Gaithersburg, MD 20899-0001

NIST

QC
100
.U56
NO.6028
1997

NISTIR 6028

Molecular Dynamics and Hydrogen Bonds in Water

Raymond D. Mountain

U.S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards
and Technology
Physical and Chemical Properties Division
Gaithersburg, MD 20899-0001

June 1997

U.S. DEPARTMENT OF COMMERCE
William M. Daley, Secretary

TECHNOLOGY ADMINISTRATION
Gary Bachula, Acting Under Secretary for Technology

NATIONAL INSTITUTE OF STANDARDS
AND TECHNOLOGY
Robert E. Hebner, Acting Director

Molecular Dynamics and Hydrogen Bonds in Water

Raymond D. Mountain
Physical and Chemical Properties Division
National Institute of Standards and Technology, Gaithersburg, MD 20899, USA

ABSTRACT

This report describes the results of an extensive set of molecular dynamics simulations of the SPC/E water model over a range of densities from greater than ambient at 1200 kg/m^3 to states near the critical density at 300 kg/m^3 and over a range of temperatures from 297 K to 823 K. The emphasis is on hydrogen bonding measures and how these measures vary with density and temperature. The mean number of bonds per molecule, the distribution of the lifetimes of the bonds and the size of hydrogen bonded clusters are reported.

1. Introduction

This report contains the results of a model study of hydrogen bonding in water over a wide range of temperatures and densities. The model was examined using molecular dynamics simulations. This study was designed to examine the changes that occur in the hydrogen bonding of water molecules as the temperature and the density of the fluid are varied. The emphasis here is on water at elevated temperatures and at densities that vary from greater than ambient, 1200 kg/m^3, to ambient, 1000 kg/m^3, down to 300 kg/m^3 just below the critical density since significant changes in hydrogen bonding occur in the model over this range of thermodynamic states. These states are typically high pressure states. Molecular dynamics simulations[1] provide a means for examining features of liquids under conditions that are difficult to realize in the laboratory and/or that are not directly accessible to experimental probes. The simulation work reported here illustrates both of these aspects of simulations.

The simulations were performed using the SPC/E model[2] for water-water interactions that is described in detail in Section 2. This three site model provides property predictions that are in qualitative accord with experimental results over the range of interest. In particular, the liquid side of the liquid-vapor coexistence curve is determined with surprising accuracy as are the values of the critical temperature and density.[3,4] This gives us some confidence that the calculated trends in hydrogen bonding noted with changes in density and temperature will reflect those that occur in liquid water. The correspondence with water will be imperfect, but the overall "picture" should be reliable, at least for densities greater than the critical density.

Section 2 contains a description of the technical details of the simulations, including the "definition" of a hydrogen bond. The SPC/E model is described in detail along with the simulation methodology. Section 3 contains a list of the thermodynamic states examined and tables of computed quantities. The pressure, internal energy, self-diffusion coefficients, the number of hydrogen bonds per molecule, and the average time required for a single molecule dipole moment to reorient are listed. Some other quantities that shed additional light on the changes that occur in the hydrogen bonding as the density and temperature vary are also mentioned. In particular, the distribution of

the lifetimes of hydrogen bonds, the size distribution of hydrogen bonded clusters, and an observable quantity that correlates with the change in the size of hydrogen bonded clusters are described. Finally, Section 4 contains a discussion of the results of the simulation.

2. The SPC/E model and technical details of the simulations

The simulations are performed on systems containing 216 water molecules located in a cubic cell. Periodic boundary conditions are applied in all three dimensions. These molecules interact via the SPC/E model potential.[2] This model contains three rigidly constrained interaction sites. The "oxygen" site is separated from two "hydrogen" sites by a distance of 0.1 nm. The HOH angle is 109.47°. An electric charge of -0.8476e is located at the oxygen site and compensating charges of 0.4238e are located at the hydrogen sites. All intermolecular sites interact through the Coulomb interaction. Also the intermolecular oxygen sites interact through a Lennard-Jones potential with ϵ/k_B = 78.24 K and σ = 0.316 nm. The potential energy, U, is

$$U = \sum_{i<j} 4\epsilon\Big[\big(\sigma/r_{ij}\big)^{12} - \big(\sigma/r_{ij}\big)^{6}\Big] + \sum_{i<j}\sum_{\alpha,\beta} q_{i,\alpha}q_{j,\beta}/r_{i\alpha,j\beta} + E_{ind}. \qquad (1)$$

The induction energy associated with the excess dipole moment is E_{ind} = 5.213 kJ/mol. In the first term, the Lennard-Jones sum, r_{ij} is the oxygen-oxygen separation and in the second term, the Coulomb sum, $r_{i\alpha,j\beta}$ is the distance between site α on molecule i and site β on molecule j. This model contains five parameters. These are the bond length and the angular separation of the hydrogen sites, the two Lennard-Jones parameters, and the charge on the oxygen site. While one frequently refers to the interaction sites as atomic sites, it must be realized that the two types of sites are not identical. A molecule has an effective dipole moment $\mu = 7.84 \times 10^{-30}$ Cm (2.35 D) that is larger than the dipole moment of the isolated molecule, 6.17×10^{-30}Cm (1.85 D). This enhanced effective moment is a mean-field way of incorporating into the model the induced polarization of a water molecule due to the field generated by neighboring molecules.[5] This is the source of the induction energy, E_{ind}. It is expected that this mean-field approach overemphasizes the induction effect as the density of the fluid decreases from the ambient density. At present, it is not known how to refine this statement. The SPC/E model is known to slightly underestimate the dielectric constant of the fluid at ambient conditions.[6] Experience with water models indicates that the effective moment should not be larger than 8.76×10^{-30} Cm (2.6 D) if "plausible" values of the dielectric constant are to be provided by a water model.[7] It appears that dielectric constant is not a useful probe of changes in the induced moments since the predicted dielectric constant is not overly sensitive to the explicit value of the effective moment. Also, the temperature-density range where changes in the induction are significant is probably where the dielectric constant is small. Some other probe will be needed to address this interesting topic.

The long range Coulomb interactions are treated using the Ewald summation method.[8] The adaptation of this method to rigid, extended molecules developed by Nosé and Klein[9] is used to determine the value of the pressure. The Lennard-Jones interactions are truncated at one-half the simulation cell edge and the quoted values for the energy and the pressure include a long range correction.[10] The orientational degrees of freedom are described using quaternions.[11] The equations of motion for the system are integrated using an iterated form of the Beeman algorithm[12] with a time step Δt = 1 fs. The duration of each simulation is 100 ps after an equilibration run of at least 10 ps.

The volume of the system is held fixed for a particular density and temperature. The temperature of the system is maintained at the desired value using separate Nosé-Hoover thermostats for the translational and rotational degrees of freedom.[13] The equations of motion for this type of system

are discussed elsewhere.[14] The inertial terms that control the coupling between the the thermostats and these degrees of freedom have a value of $Q = 0.05m\sigma^2$ where m is the mass of a water molecule. This value provides a reasonably steady value for the translational and rotational kinetic energies. Thus, averages are determined in the NVT or canonical ensemble.

Figure 1 summarizes the thermodynamic states that we examine in this study and these states will be discussed in some detail below. In order to place these states in a sensible context, the liquid-vapor coexistence curves for water[15] and for the SPC/E system determined by Guillot and Guissani[3] are displayed.

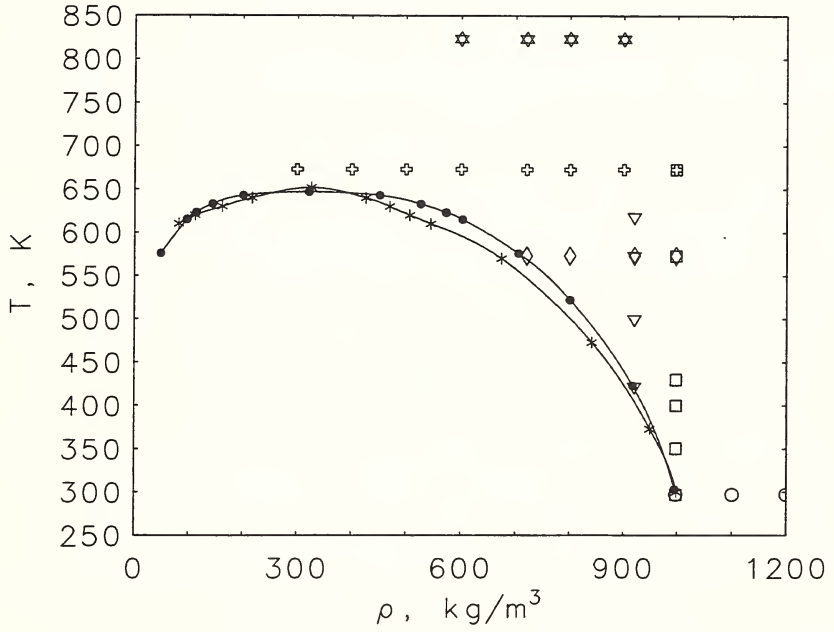


Figure 1. This is the ρ-T plane for SPC/E water showing the state points that have been simulated. The solid line through the solid dots is the liquid-vapor coexistence curve for water. The solid line through the asterisks (*) indicates the liquid-vapor coexistence curve for the SPC/E model. The other symbols indicate various isotherms and isochores that are discussed in the text.

A number of quantities are monitored during each simulation. These quantities will be introduced in the following paragraphs along with an exposition of the utilization of these quantities in Sections 3 and 4. It is illustrated using results for ambient water (T = 297 K and ρ =997 kg/m^3).

The first quantity is the energy that is described by eq. 1. Next, the pressure is obtained using the molecular virial formulation,[16,17]

$$pV = Nk_BT + \frac{1}{3}\langle \sum_{i<j} \mathbf{R}_{ij} \cdot \mathbf{F}_{ij} \rangle \qquad (2)$$

where $\mathbf{R}_{ij}$ and $\mathbf{F}_{ij}$ are the distance between molecular centers of mass and the force between molecules i and j. The $\langle \ldots \rangle$ notation indicates an ensemble average of the enclosed quantity. For our purposes, ensemble averages and time averages are equivalent.[18] Also the intermolecular site-site pair correlation functions $g_{OO}(r)$, $g_{OH}(r)$, and $g_{HH}(r)$ are determined. The oxygen-oxygen pair function, $g_{OO}(r)$, is used to determine size of the neighborhood around a given molecule that contains four near neighbor molecules. To determine this size, we first construct the coordination number, N(R), as a function of the distance from a given molecule;

$$N(R) = 4\pi n \int_0^R dr r^2 g_{OO}(r), \tag{3}$$

where n is the number density of molecules in the fluid. Then we define a quantity, R_4 such that

$$N(R_4) = 4 \tag{4}$$

is satisfied, as illustrated in Fig. 2. We find that R_4 is insensitive to the temperature and at high density has a value of about 0.32 nm. It will be shown below that significant departures from this value are associated with large scale changes in local structure.

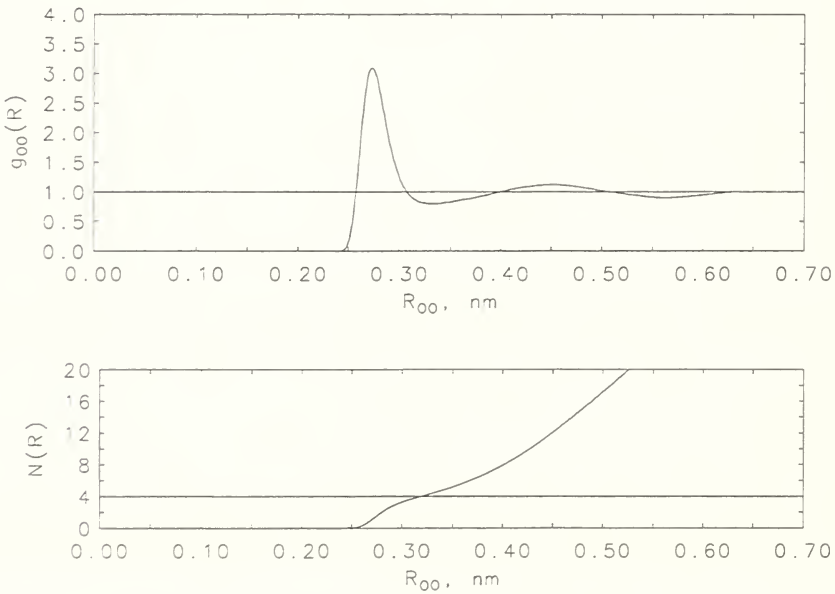


Figure 2. This illustrates the use of eqs. 3 and 4 to determine R_4 to be 0.32 nm. This example is for the state with T = 297 K and ρ = 997 kg/m^3. Note that the position of the secondary maximum in $g_{OO}(r)$ is about 1.6 times the position of the principal maximum.

The oxygen-hydrogen pair function, $g_{OH}(r)$, is used to determine the number of hydrogen bonds per molecule. A measure of the number of hydrogen bonds, N_B, is obtained from an integral over the first maximum in g_{OH} that is analogous to the one in eq. 3. This maximum is referenced as the hydrogen bond signature in this article. This integral is

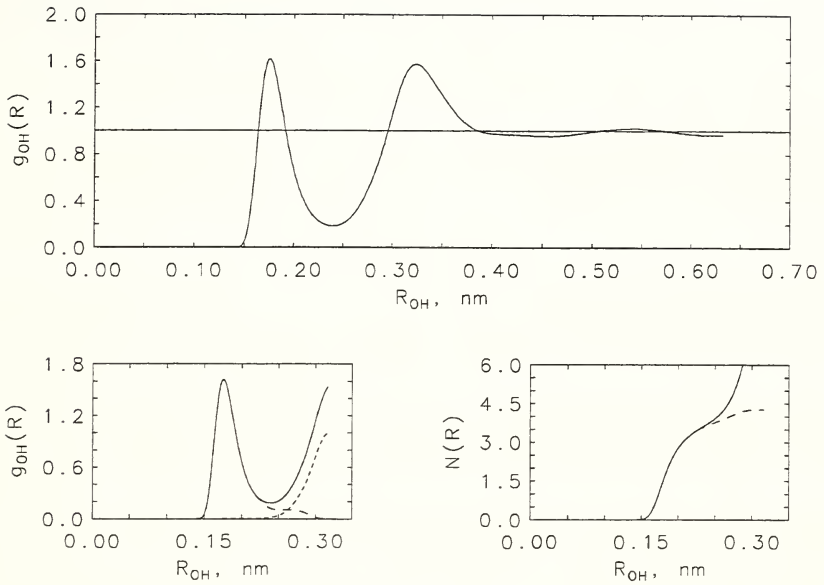


Figure 3. This illustrates the use of eq. 5 to determine the number of hydrogen bonds, N_B, for the state with T = 297 K and ρ = 997 kg/m^3. The upper plot is g_{OH}(r), the lower left plot shows the hydrogen bond signature in more detail. Also shown is the distribution of pairs that satisfy the geometrical criterion (dashed line) and the difference between the two curves (short dashed line). The lower right plot shows the coordination numbers obtained using the two criteria.

$$N_B = 4\pi n_B \int_0^{R_{min}} dr r^2 g_{\mathrm{OH}}(r), \tag{5}$$

where n_B is the number density of the fluid times the number of hydrogen bonds that a given molecule can experience, so n_B = 4n. The quantity, R_{min}, is the position of the first minimum in g_{OH} and is found at approximately 0.24 nm. This experimentally accessible measure of the number of hydrogen bonds has been shown within the framework of a molecular dynamics simulation[19] to yield values for N_B that are consistent with those obtained using an explicit geometrical criterion described by Mezei and Beveridge.[20] This geometrical criterion states that two water molecules are hydrogen bonded if the O-O separation is less than 0.33 nm, the OH intramolecular bond on the donor molecule makes an angle with the O-O axis that is less than 70.53°, and that the acceptor molecule be similarly oriented in terms of acceptor sites that are tetrahedrally related to the intramolecular hydrogen sites. At low temperatures, the two measures agree rather closely as may seen by comparing the solid and dashed lines in the lower right plot Figure 3. (The short dashed line is the difference.) However, at elevated temperatures, the first measure is systematically greater than the geometrical one.[21]

Another hydrogen bond property is the lifetime of a specific bond. Each hydrogen site is monitored during the simulation and the distribution of the duration of hydrogen bonds, $p_B(t)$, is determined with a time resolution of 0.1 ps. Here the geometrical condition is used to define a hydrogen bond. At the same time, the duration of the nonbonded condition is also determined. The results of these lifetime distributions are used to determine the cumulative lifetime distribution, $f_N(t)$;

$$f_N(t) = C \int_0^t dssp_B(s). \tag{6}$$

The corresponding distribution for the lifetime of the nonbonded condition is

$$f_{NB}(t) = C \int_0^t dssp_{NB}(s) \tag{7}$$

and the constant C is detemined by the condition

$$f_N(25) + f_{NB}(25) = 2. \tag{8}$$

For all states examined in this study, no bonds persisted for as long as 25 ps.

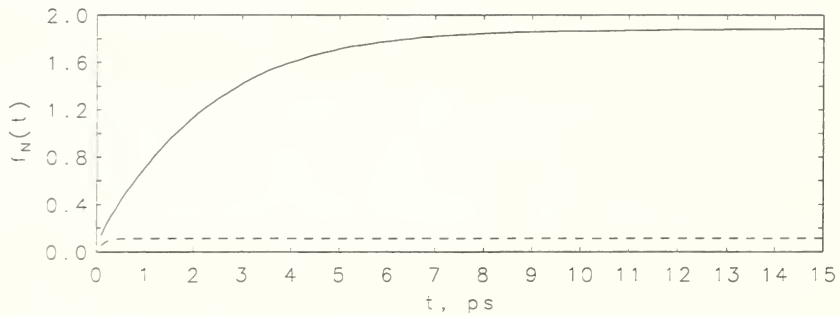


Figure 4. The solid curve is the cumulative distribution of hydrogen bond lifetimes and the dashed curve is distribution of lifetimes for the nonbonded condition for hydrogen sites at the state with T = 297 K and ρ = 997 kg/m^3.

At the same time as the lifetime information is collected, a second quantity is determined. This is the size distribution of hydrogen bonded clusters. Such a cluster is defined as follows.[22] If two molecules are hydrogen bonded, they are members of a specific hydrogen bonded cluster. Given the fact that the state of hydrogen bonding is known a some instant, it is a straightforward sorting process to determine the number of distinct hydrogen bonded clusters and how many molecules are in each cluster. A FORTRAN77 subroutine, `cluster`, that is listed in Appendix A, provides the information needed to generate $\mathcal{N}$(N), the mean number of clusters containing N molecules. A more useful quantity is N$\mathcal{N}$(N) which is proportional to the average number of molecules contained in clusters of size N. It is convenient to convert this quantity into a density by normalizing it so that

$$\int_0^{N_{total}} dxx\mathcal{N}(x) = 1. \tag{9}$$

With this definition, the mean number of molecules in a cluster is

$$\langle n \rangle = \int_0^{N_{total}} dx x^2 \mathcal{N}(x). \tag{10}$$

Thus, the interesting quantity is the mass weighted distribution of cluster size as a function of f =N/N_{total}, the fraction of the total number of molecules in a cluster,

$$P_c(f) = n^2 \mathcal{N}(n)/N_{total}. \tag{11}$$

The analysis of cluster sizes is made every 0.1 ps during each simulation. An example is shown in Figure 5.

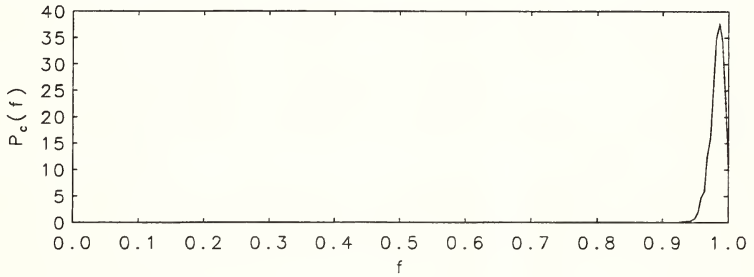


Figure 5. The mass weighted distribution of cluster size, P_C(f), in a cluster is shown as a function of f, the fraction of the total sample that is in a cluster for the state with T = 297 K and ρ = 997 kg/m^3. The small number of molecules not in the main cluster does not register on this scale.

Two other dynamical quantities are determined during each simulation. The first is $R^2(t)$, the mean square displacement of a molecule as a function of time,

$$R^2(t) = \langle \frac{1}{N} \sum_{i=1}^{N} \mid \mathbf{R}_i(t) - \mathbf{R}_i(0) \mid^2 \rangle. \tag{12}$$

The self-diffusion coefficient, D_S, is obtained from R^2(t) by the Einstein relation[23]

$$D_S = \lim_{t \to \infty} \frac{1}{6} \frac{dR^2(t)}{dt}. \tag{13}$$

In practice, the long time limit is obtained within 1-2ps. The second dynamical quantity is the time for an individual dipole moment to reorient. It is obtained by constructing the following time autocorrelation function. Let $\hat{\mathbf{d}}_i$ be a unit vector lying along the dipole moment of molecule i, and let

$$C(t) = \frac{1}{N} \sum_i \langle \hat{\mathbf{d}}_i(t) \cdot \hat{\mathbf{d}}_i(0) \rangle. \tag{14}$$

The mean time for a dipole to reorient, t_1 satisfies

$$C(t_1) = C(0)\exp(-1). \tag{15}$$

3. Results

Tables 1 and 2 contain the computed values of several of the quantities discussed in the previous section. The organization of the tables follows the sequence of states displayed in Figure 1. It starts with the highest density, lowest temperature state and moves to higher temperatures and lower densities. For a given density, the states with increasing temperature are listed in order. Table 1 contains values for the potential energy as described by eq. 1, the pressure as described by eq. 2, and the self-diffusion coefficient as developed using eqs. 12 and 13. The pressure and the energy values include the contribution of the long range part of the Lennard-Jones terms that are neglected in the force calculation.[10] While no error analysis has been attempted for these quantities, the accuracy of these estimates is commensurate with the number of digits listed in Table 1 and in Table 2.

Table 1. The potential energy, pressure, and self-diffusion coefficients for the states examined in this study.

T,K	ρ_m, kg/m^3	U, kJ/mol	p, MPa	D_S, 10^{-9} m^2/s
297	1200.	-40.7	726.	1.84
297	1100.	-42.1	260.	2.11
297	997.	-41.6	-15.	2.23
350	997.	-38.9	65.	5.28
400	997.	-36.6	145.	8.73
430	997.	-35.3	206.	10.4
573	997.	-30.6	503.	17.9
673	997.	-28.0	715.	21.8
422	920.	-34.7	28.	11.9
500	920.	-31.7	158.	17.3
573	920.	-29.5	286.	20.7
618	920.	-28.2	365.	23.9
673	900.	-26.6	418.	27.3
823	900.	-23.3	671.	33.7
573	800.	-27.5	92.	27.8
673	800.	-24.7	218.	33.0
823	800.	-21.5	427.	38.7
573	720.	-25.9	24.	31.3
673	720.	-23.1	129.	39.4
823	720.	-20.0	294.	46.6
673	600.	-20.7	59.	50.2
823	600.	-17.4	174.	58.1
673	500.	-18.3	34.	60.3
673	400.	-16.1	25.	71.5
673	300.	-13.1	23.	86.9

Table 2. The mean number of hydrogen bonds per molecule, N_B, R_4, and t_1, the reorientation time for an individual dipole for the states examined in this study.

T,K	ρ_m, kg/m^3	N_B	R_4, nm	t_1, ps
297	1200.	3.9	0.30	3.4
297	1100.	3.7	0.31	3.8
297	997.	3.6	0.32	4.3
350	997.	3.5	0.32	1.8
400	997.	3.4	0.31	1.0
430	997.	3.3	0.32	0.79
573	997.	3.0	0.32	0.33
673	997.	2.9	0.32	0.23
422	920.	3.3	0.32	0.88
500	920.	3.0	0.32	0.48
573	920.	2.8	0.33	0.33
618	920.	2.8	0.33	0.26
673	900.	2.5	0.33	0.22
823	900.	2.2	0.33	0.13
573	800.	2.6	0.34	0.29
673	800.	2.3	0.34	0.21
823	800.	2.0	0.34	0.14
573	720.	2.5	0.35	0.33
673	720.	2.2	0.35	0.21
823	720.	1.9	0.35	0.12
673	600.	2.0	0.37	0.19
823	600.	1.6	0.37	0.09
673	500.	1.6	0.39	0.16
673	400.	1.4	0.41	0.16
673	300.	1.2	0.44	0.12

Table 2 contains the estimates for three other derived quantites. These are N_B, the mean number of hydrogen bonds per molecule, as determined using eq. 5, the R_4 parameter determined using eqs. 3 and 4, and the mean dipole reorientation time determined using eqs. 14 and 15.

The other quantities described in Section 2, the site-site pair functions, the distribution of hydrogen bond lifetimes, and the hydrogen bonded cluster size distribution will be discussed in the following section.

4. Discussion

There are four distinct regions of the temperature-density plane to be discussed. As is done in the tables, these regions are arranged by decreasing density and increasing temperature. In each region we examine the number of hydrogen bonds/molecule, N_B, the distribution of the lifetimes of hydrogen bonds, $f_B(t)$, the mass weighted size distribution of hydrogen bonded clusters, $P_c(n)$, the time required for reorientation of individual dipoles, t_1, the nature of the short-range order in the liquid, and the self-diffusion coefficient D_S.

Region 1

First, we examine the low temperature, high density region. This is the set of states on the 297 K isotherm over the density interval from 997 kg/m^3 to 1200 kg/m^3. The striking feature of this region is the crowding together of molecules as the density is increased. The shoulder that develops on the outer side of the principal maximum of $g_{OO}(r)$ as shown in Figure 6 is indicative of this. Note from Table 2 that the value of R_4 remains roughly constant suggesting that the inherent 4-fold coordination of the water molecules is maintained. This sort of behavior was noted by Madura, *et al.*[24] for a different potential model (TIP4P)[25].

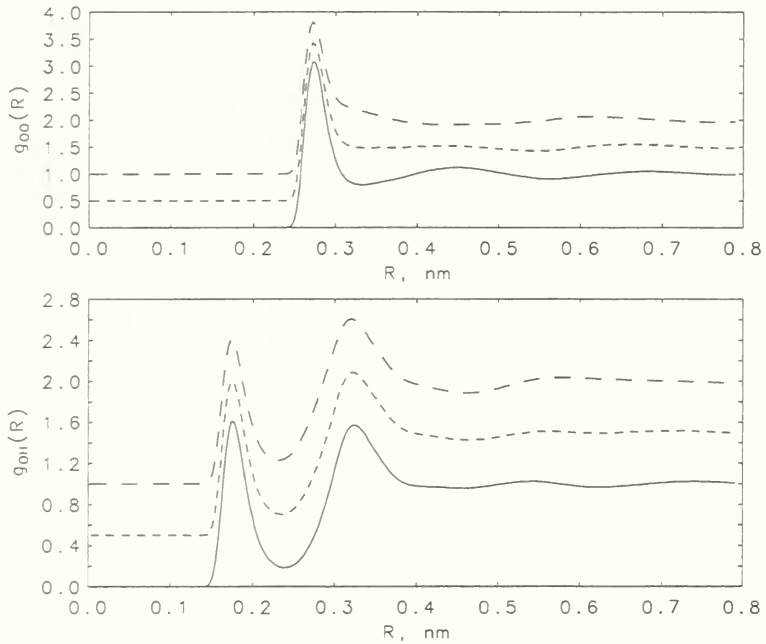


Figure 6. The oxygen-oxygen and oxygen-hydrogen pair functions are displayed for the 297 K isotherm. The solid line is the 997 kg/m^3 state, the short-dash line is for the 1100 kg/m^3 state and the long-dash line is for the 1200 kg/m^3 state. The curves are offset by 0.5.

Again referring to Table 2, we see that the dipole reorientation time, t_1, decreases with increasing density while the mean number of hydrogen bonds increases. This seems surprising at first so it is instructive to characterize the bonding in more detail. The hydrogen bond signature of $g_{OH}(r)$ in Figure 6 and the distribution of hydrogen bond lifetimes in Figure 7 provide a rationalization of these changes in N_B and t_1. The height of the first maximum in $g_{OH}(r)$ decreases with increasing density so the increase in N_B is due to the increase in density. Also, we see in Figure 7 that the

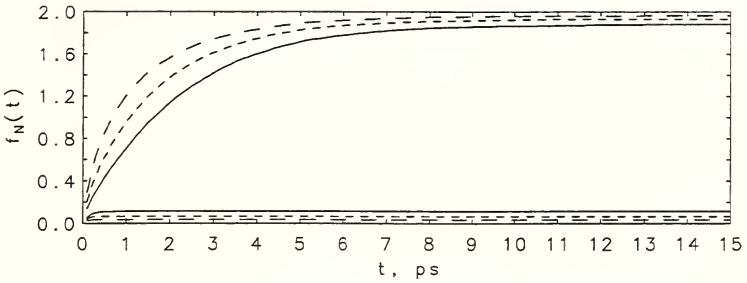


Figure 7. The cumulative distributions for the lifetimes of hydrogen bonds (upper set) and for the lifetimes of the unbonded condition for the hydrogen sites (lower set) for the 297 K isotherm. The solid line is the 997 kg/m^3 state, the short-dash line is for the 1100 kg/m^3 state and the long-dash line is for the 1200 kg/m^3 state.

fraction of long-lived hydrogen bonds decreases with increasing density. This trend is consistent with the corresponding decrease in t_1.

Region 2

The second region examined is the 297 K to 673 K portion of 997 kg/m^3 isochore. This is a region where significant changes in hydrogen bonding occur. The number of bonds/molecule decreases smoothly, the self-diffusion coefficient increases rapidly and the dipole reorientation time decreases rapidly, as shown in Figure 8.

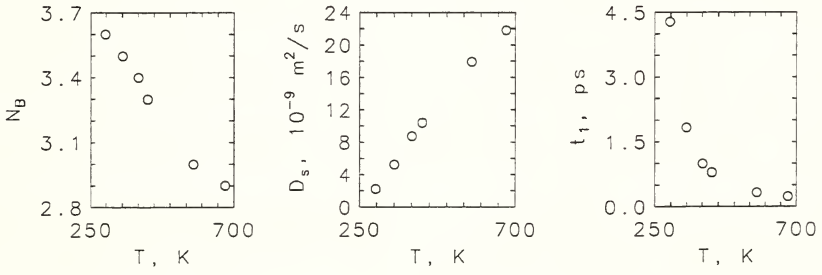


Figure 8. The temperature dependence of N_B, D_S, and t_1 along the 297 kg/m^3 isochore from 297 K to 673 K is shown here.

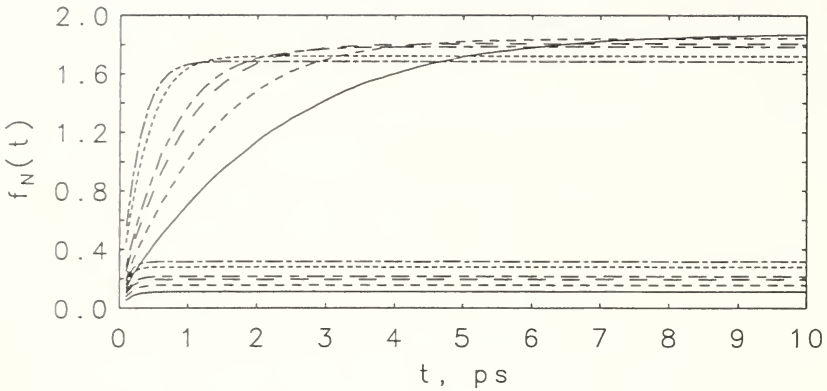


Figure 9. The temperature dependence of the distribution of the lifetimes of hydrogen bonds on the the 997 kg/m^3 isochore from 297 K to 673 K is shown here. The lower set of curves is the distribution of the lifetimes on the nonbonded condition for hydrogen sites. The lower set increase upward in temperature starting with the solid line for 297 K, and moving upward through 350 K, 400 K, 430 K, 573 K, to the long-short dashed line for 673 K.

More striking is the change in the distribution of the lifetimes of the bonds shown in Figure 9. As the temperature increases, the shape of this curve changes fairly rapidly between 400 K and 573 K. It is in this region that the maximum lifetime of a bond drops to about 1 ps.

In this same temperature interval the short range order undergoes a significant change, as illustrated in Figure 10. The decrease in the number of hydrogen bonds is associated with the decrease in the height of the first maximum in $g_{OH}(r)$. This feature also becomes effectively broader as the temperature increases. The position of the second maximum in $g_{OO}(r)$ shifts from 0.45 nm to 0.6 nm. This change indicates that the tetrahedral coordination is no longer persists to the second neighbor distance. R_4 remains unchanged so the first neighbor coordination is maintained.

It should be noted that over this temperature interval the mass weighted distribution of cluster sizes, $P_c(n)$, remains concentrated at the high end of the size scale. The majority of the molecules continue to be in one large hydrogen bonded cluster.

Let us summarize the changes in hydrogen bonding that occur in Region 2. The number of hydrogen bonds decreases and the lifetime of the bonds decreases markedly, particularly at the higher temperature range. Also, the reorientation time for the individual dipoles decreases rapidly, possibly as a consequence of the reduced constraints on the rotation of the molecules. These changes also show up in the oxygen-oxygen pair function as a shift in the position of the secondary maximum from about 0.45 nm to about 0.6 nm. This is generally viewed as a change in the distance over which tetrahedral structure is correlated.

The changes in the pair functions can be roughly compared with the corresponding experimental quantities.[26] Curves 1 and 2 of Figs. 2-4 in Ref. 26 are neutron scattering based estimates of the

site-site pair functions at 997 kg/m^3 and at 298 K and 423 K respectively. There is qualitative agreement between the experimental pair functions and those shown in Figure 10. In particular, the change in the second maximum in $g_{OO}(r)$ and the weakening of the hydrogen bonding signature in $g_{OH}(r)$ with increasing temperature found in the simulation results are present in the experimental pair functions.

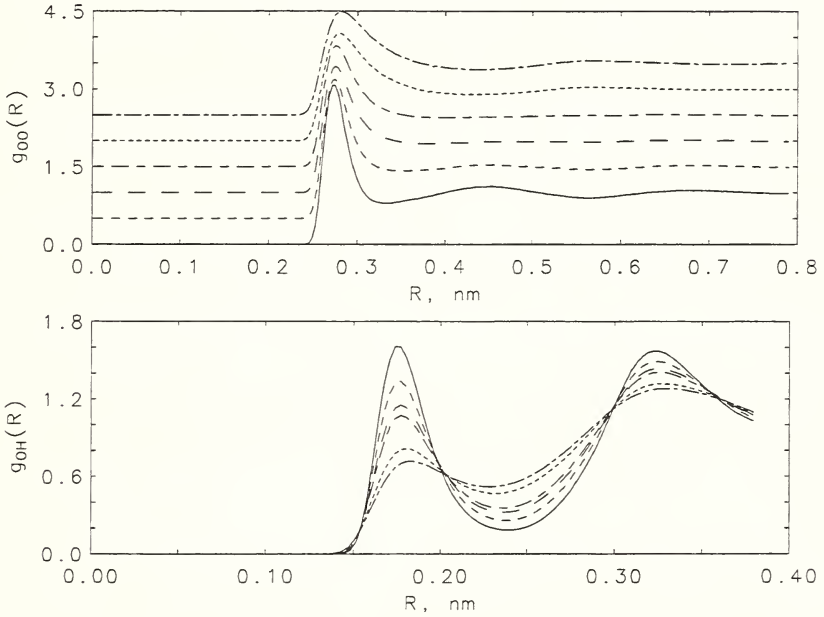


Figure 10. The temperature variation of the oxygen-oxygen and oxygen-hydrogen pair functions along the 997 kg/m^3 isochore from 297 K to 673 K is illustrated here. The oxygen-oxygen functions are offset by 0.5 for clarity starting with the solid line for 297 K, and moving upward through 350 K, 400 K, 430 K, 573 K, to the long-short dashed line for 673 K. The oxygen-hydrogen functions are superimposed to emphasize the strong variation with temperature of the hydrogen bond signature.

Region 3

The third region contains the high temperature isotherms, 573 K and 673 K over the density interval 997 kg/m^3 down to 600 kg/m^3. Figure 11 shows the variation of the number of hydrogen bonds and the variation of R_4 with decreasing density. The drop in the number of hydrogen bonds with decreasing density matches the abrupt rise in R_4. The change in R_4 is not reflected in the self-diffusion coefficient. This is probably due to the different time scales for reorientation of a dipole, about 0.3 ps, the lifetime of hydrogen bonds, less than 1 ps and the time of 1 ps for significant

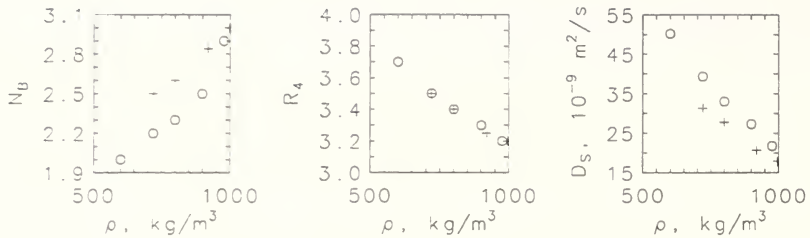


Figure 11. The number of hydrogen bonds, N_B, the quantity R_4, and the self-diffusion coefficient, D_S for the 573 K isotherm (+'s) and for the 673 K isotherm (circles) between 997 kg/m^3 and 600 kg/m^3.

displacement of a molecular diameter $\approx$ 0.3 nm. The environment in which a molecule diffuses is averaged by the rapid changes in molecular orientation at these temperatures.

The increase in R_4 suggests that changes are occuring in the short range correlation of the molecules. The pair functions shown in Figure 12 contain a hint. The hydrogen bond signature that is rapidly decreasing in Region 2 is nearly constant in Region 3. If anything, the amplitude of the maximum increases slightly. Any changes in the number of hydrogen bonds per molecule are due to density changes rather than short range structural changes.

However, if one takes a larger scale view, something interesting emerges. The evolution of this distribution of cluster sizes, P_c, along the 673 K isotherm as the density is varied from 997 kg/m^3 to 600 kg/m^3 is shown in Figure 13. Here we see that the hydrogen bonded cluster size drops rapidly with decreasing density at this temperature. Instead of basically one cluster with a few non-bonded molecules found at high density, the distribution of cluster sizes becomes very broad with no single size dominating. That this change correlates with the changes in R_4 is quite interesting. The transition from a hydrogen bond network that includes most of the molecules to a collection of much smaller, disjoint clusters, has an *experimentally observable consequence*.

It is possible to make contact with experiments in this region as well. Reference 26 contains site-site pair functions for states in this region. Curves 4–9 in Figs. 3 and 4 of that reference can be compared with the simulation results shown in Figure 12. There is qualitative agreement with the neutron scattering based pair functions for densities greater than 700 kg/m^3. A serious discrepancy between the simulations and the experimental $g_{OH}(r)$ occurs for lower densities. The experimentally derived function has the hydrogen bond signature reduced to a shoulder while the simulation function is insensitive to changes in the density. The experimental result suggests that a major breakdown of hydrogen bonding occurs over a narrow density interval while the model results do not indicate such a change.

Another experimental measure of hydrogen bonding is the proton NMR chemical shift. Recent work[27,28] indicates that N_B at 673 K and 520 kg/m^3 is on the order of 1. This result contradicts the neutron scattering result that there is little hydrogen bonding. A simulation based estimate of the chemical shift at low temperatures overestimated the shift by 15–20% for the SPC/E model.[29]

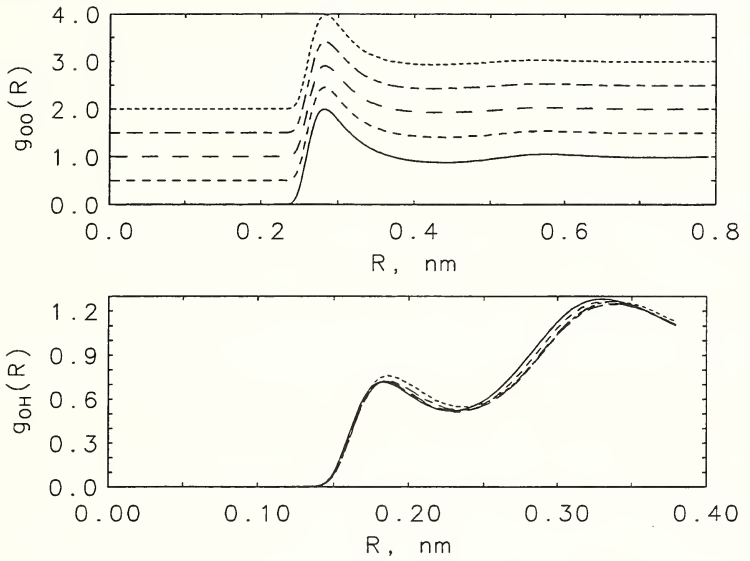


Figure 12. Oxygen-oxygen and oxygen-hydrogen pair functions for the 673 K isotherm. The solid line is the 997 kg/m^3 state. The $g_{oo}(r)$'s have been offset upward by 0.5 for the successive densities, 900 kg/m^3, 800 kg/m^3, 720 kg/m^3, and 600 kg/m^3. The hydrogen-oxygen pair functions are superimposed to emphasize the insensitivity of the hydrogen bond signature to changes in the density.

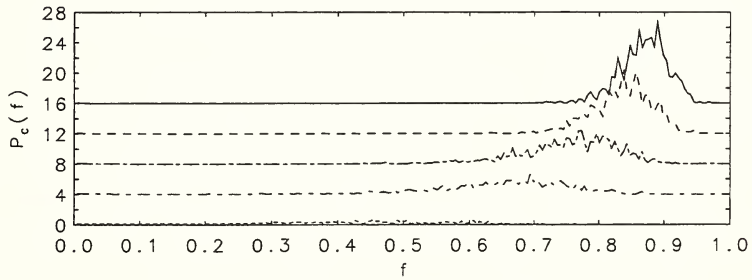


Figure 13. The cluster size distributions, $P_c(f)$ are shown the 673 K isotherm. The solid line is the 997 kg/m^3 state. The P_c's have been offset by 4 units between successive densities. Moving down from the 997 kg/m^3 curve, the curves are for 900 kg/m^3, 800 kg/m^3, 720 kg/m^3, and 600 kg/m^3.

The neutron scattering results indicate that the mean-field model has reached a dramatic end to the limit of applicability, at least as far as being able to describe hydrogen bonding. The chemical shift results instead imply that the limit has not been reached.

Finally, the situation at the low density end of this region is that the lifetimes of hydrogen bonds and of nonbonded states for the hydrogens have become equal. This is shown in Figure 14 .

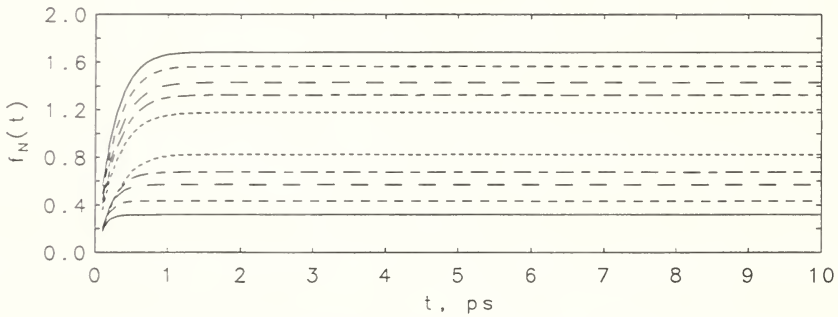


Figure 14. The temperature dependence of the distribution of the lifetimes of hydrogen bonds on the the 673 K isotherm for densities of 997 kg/m^3 (solid line), decreasing through 900 Kg/m^3, 800 Kg/m^3, 720 Kg/m^3, to 600 Kg/m^3. The lower set of curves is the distribution of the lifetimes on the nonbonded condition for hydrogens.

Region 4

The final region to be examined is the continuation of the 673 K isotherm down to lower density, 300 kg/m^3. Here the hydrogen bonded clusters are small and contain only a small fraction of the total number of molecules, as displayed in Figure 15. There, $\langle f \rangle = \langle n \rangle / N_{total}$ is shown for both Regions 3 and 4. Equation 10 was used to determine $\langle n \rangle$.

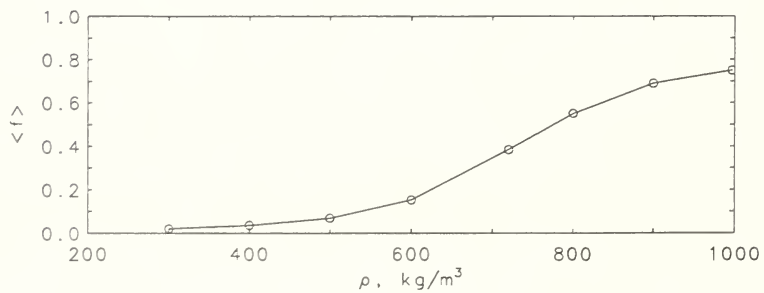


Figure 15. This is the average fraction of the total number of molecules in a cluster along the 673 K isochore. The solid line is a guide to the eye.

In this region the lifetime of the nonbonded hydrogen is longer than that of the hydrogen bond, as shown in Figure 16.

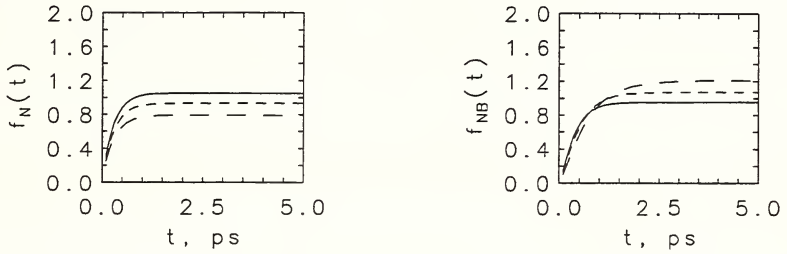


Figure 16. The cumulative distributions of lifetimes for hydrogen bonds and for the nonbonded state are shown here for Region 4; solid line, 500 kg/m^3, short dash line, 400 kg/m^3, long dash line 300 kg/m^3.

Another interesting development is how the oxygen-oxygen pair function changes as the density decreases from 600 kg/m^3 to 300 kg/m^3. From Table 2 we see that the value of R_4 *increases* as the density decreases. In Figure 17, the oxygen-oxygen and oxygen-hydrogen pair functions are displayed. The height of the first maximum for both functions increases with decreasing density.

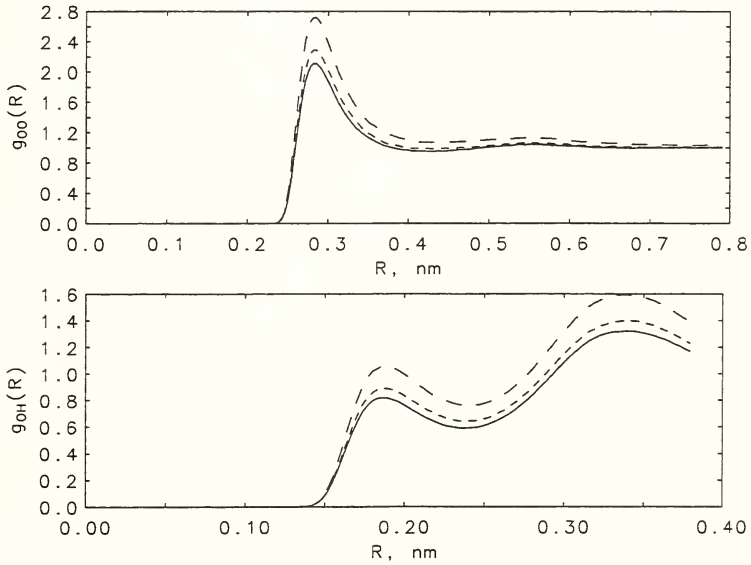


Figure 17. Oxygen-oxygen and oxygen-hydrogen pair functions for the 673 K isotherm in Region 4. The solid line is the 500 kg/m^3 state, the short dash line is for the 400 kg/m^3, and the long dash line is for the 300 kg/m^3 state.

This means that short range order is increasing at the expense of longer range structure. The first figure in Appendix B provides a visual comparison of a configuration at 600 kg/m^3 with one of a configuration at 300 kg/m^3. These are "snapshots" of configurations. It is evident that the fluid is becoming inhomogeneous over a distance spanned by a few molecules so physical clustering is developing on the molecular length scale. The increase in the first maximum in $g_{OO}(r)$ with decreasing density is a signature of this effect of physical clustering. This clustering is necessarily a transient in the sense that the fluid is homogeneous when examined over a sufficiently long time interval. The existence of clusters has been suggested on the basis of simulations[30] and from an analysis of X-ray diffraction data.[31]

Is this clustering unique to water?

To answer this question, consider the Lennard-Jones fluid. This is a monatomic fluid of particles that interact through a pair-wise additive potential

$$V(r) = 4\epsilon\left[(\sigma/r)^{12} - (\sigma/r)^{6}\right]. \tag{16}$$

Temperatures for this system are defined in units of ϵ/k_B and the densities in units of $n = \rho\sigma^3$. The critical temperature of the Lennard-Jones fluid is $T \approx 1.35$ and the critical density is $n \approx 0.33$.[10]

A system of 500 Lennard-Jones particles was simulated with the purpose of examining the configurations for signs of physical clustering. Each state was run for 100 τ time units. (The natural time unit for the Lennard-Jones system is $\tau = \sqrt{m\sigma^2/\epsilon}$ where m is the mass of the particle and ϵ and σ are the parameters in eq. 16.) Figure 18 displays the g(r) for the 1.4 isotherm for densities varying from 0.6 (solid line) down to 0.2 (long dash line). While the magnitude of change in g(r) with decreasing density is not as pronounced as in water, it is present. The inset in Figure 18 shows the height of the maximum of g(r) as a function of the density. The minimum in the height of g(r) occurs around $\rho \approx 0.35$. The snapshot view of configurations for n = 0.4 and for n = 0.2 is the second figure in Appendix B. Again, the local inhomogeneity develops as the density decreases although the effect is not as pronounced as it is for SPC/E water.

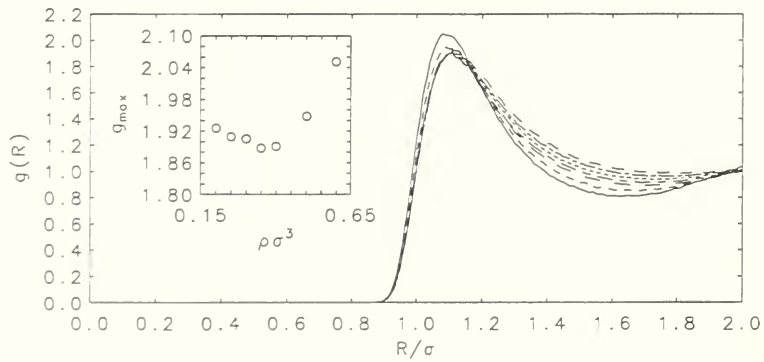


Figure 18. The pair functions for the T = 1.4 isotherm of the Lennard-Jones fluid for several densities, 0.6, 0.5, 0.4, 0.35, 0.3, 0.25, and 0.2. The solid line is for 0.6 (the deepest minimum) and the long-dash line is for 0.2 (the shallowest minimum. The inset shows the maximum peak height as a function of density.

The answer then is no, physical clustering of this sort is not unique to water. It is however, more pronounced in water, probably due to the much stronger electrostatic interactions found in water relative to the Lennard-Jones case. To emphasize this, we show some results for the Stockmayer potential fluid.[23] The Stockmayer potential is a Lennard-Jones plus dipole-dipole interaction model,

$$U = \sum_{i<j}^{N} \left[V(r_{ij}) + \left[\frac{\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j}{r_{ij}^3} - \frac{3(\boldsymbol{\mu}_i \cdot \mathbf{r}_{ij})(\boldsymbol{\mu}_j \cdot \mathbf{r}_{ij})}{r_{ij}^5} \right] \right], \tag{17}$$

where $V(r)$ is given by eq. 16 and $\boldsymbol{\mu}_i$ is the dipole moment of molecule i. The dipoles are of fixed magnitude but are free to reorient under the mutual interactions with other dipoles. It is useful to characterize the dipole in terms of the dimensionless quantity $\mu^{*2} = \mu^2/\epsilon\sigma^3$ where μ is the magnitude of the dipole moment, and ϵ and σ are the Lennard-Jones potential parameters. The liquid-vapor phase diagram for the Stockmayer fluid has been determined for several values of the dipole moment and the critical temperatures and densities are known for μ^{*2} = 1.0, 2.0., 3.0, 4.0, 5.0, and 6.0.[32–35]

Simulations for temperatures about 10% above the critical temperature were performed for values of μ^{*2} = 1.0, 2.0., 3.0, and 6.0. These simulations were for systems of 500 particles. The long-range part of the dipole-dipole interations were treated using the Ewald summation method for dipoles.[1] Each simulation was for 100 τ time units once the fluid was equilibrated. The magnitudes of the peak in g(r) for each value of μ^{*2} are shown in Fig. 19. This shows that the magnitude of the signature of physical clustering increases with increasing strength of the dipole moment.

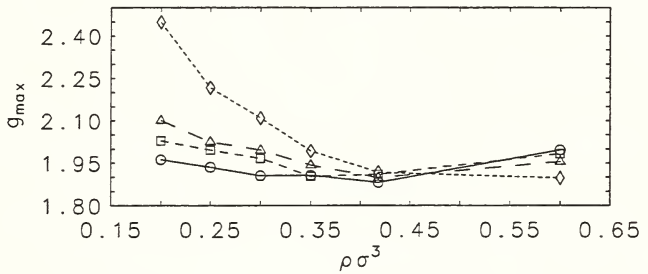


Figure 19. The maximum value of g(r) for the Stockmayer fluid along the T/T_c = 1.1 isotherm as a function of the density. The circles are for μ^{*2} = 1, the squares are for μ^{*2} = 2, the triangles are for μ^{*2} = 3 and the diamonds are for μ^{*2} = 6.

The value of μ^{*2} for the SPC/E model is 16.2 as determined using the parameters appearing in eq. 1. A quite different set of Stockmayer parameters for water have been determined.[35] With ϵ = 382.3k_B, σ =0.2955 nm, and μ^{*2} = 2.43, the liquid-vapor coexistence curve for water can be matched by the Stockmayer fluid coexistence curve. This is an example of how specific properties can be matched by a model potential. The Stockmayer fluid model is obviously incapable of describing the site specific short range order present in water.

References

1. M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, (Clarendon Press, Oxford, 1987).
2. H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, J. Phys. Chem. **91**, 6269 (1987).
3. B. Guillot and Y. Guissani, J. Chem. Phys. **99**, 8075 (1993).
4. J. Alejandre, D. J. Tildesley, and G. A. Chapela, J. Chem. Phys. **102**, 4574 (1995).
5. K. Watanabe and M. L. Klein, Chem. Phys. **131**, 157 (1989).
6. R. D. Mountain and A. Wallqvist, NISTIR 5788, "A Collection of Results for the SPC/E Water Model", 1996. This report is available from the National Technical Information Service, Springfield, VA 22161; Order Number PB96-147899/AS.
7. M. Sprik, J. Chem. Phys. **95**, 6762 (1991).
8. M. J. L. Sangster and M. Dixon, Adv. Phys. **25**, 247 (1976).
9. S. Nosé and M. L. Klein, Mol. Phys. **50**, 1055 (1983). See also, D. Brown and S. Neyertz, Mol. Phys. **84**, 577 (1995), and Ref. 4.
10. L. Verlet, Phys. Rev. **159**, 98 (1967).
11. R. Sonnenschein. J. Computational Phys. **59**, 347 (1985) and D. C. Rapaport, J. Computational Phys. **60**, 306 (1985).
12. P. Schofield, Comput. Phys. Comm. **5**, 130 (1976). The iterated form of the algorithm is described in R. D. Mountain and A. C. Brown, J. Chem. Phys. **82**, 4236 (1985).
13. G. J. Martyna, M. L. Klein, and M. Tuckerman, J. Chem. Phys. **97**, 2635 (1992).
14. R. D. Mountain, J. Chem. Phys. **105**, 10496 (1996). See eqs. 3 and 4.
15. L. Haar, J. S. Gallagher, and G. S. Kell, *NBS/NRC Steam Tables*, (Hemisphere Publishing Corporation, NY, 1984).
16. S. T. Cui, P. T. Cummings, and H. D. Cochran, Mol. Phys. **88**, 1657 (1996).
17. S. Balasubramanian, C. J. Mundy, and M. L. Klein, J. Chem. Phys. **105**, 11190 (1996).
18. D. Thirumalai and R. D. Mountain, Phys. Rev. A **42**, 4574 (1990).
19. R. D. Mountain, J. Chem. Phys. **103**, 3084 (1995).
20. M. Mezei and D. L. Beveridge, J. Chem. Phys. **74**, 622 (1981).
21. A. A. Chialvo and P. T. Cummings, J. Phys. Chem. **100**, 1309 (1996).
22. A. Geiger, F. H. Stillinger, and A. Rahman, J. Chem. Phys. **70**, 4185 (1979).
23. J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, (Academic Press, NY, 1986).
24. J. D. Madura, B. M. Pettit, and D. F. Calef, Mol. Phys. **64**, 325 (1988).
25. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983). Table 1 contains the parameters for several water potentials.

26. A. K. Soper, F. Bruni, and M. A. Ricci, J. Chem. Phys. **106**, 247 (1997).

27. N. Matubayasi, C. Wakai, and M. Nakahara, Phys. Rev. Letters **78**, 2573 (1997).

28. M.Hoffmann and M. S. Conradi, J. Am. Chem.Soc. **119**, 3811 (1997).

29. I. M. Svishchev and P. G. Kusilik, J. Am. Chem. Soc. **115**, 8270 (1993).

30. Y. Guissani and B. Guillot, J. Chem. Phys. **98**, 8221 (1993).

31. M. Nakahara, T. Yamaguchi, and H. Ohtaki, Recent. Res. Devel. Phys. Chem. **1**, 17 (1997).

32. B. Smit, C. P. Williams, E. M. Hendriks, and S. W. de Leeuw, Mol. Phys. **68**, 765 (1989).

33. M. E. van Leeuwen, B. Smit, and E. M. Hendriks, Mol. Phys. **78**, 271 (1993).

34. M. E. van Leeuwen, Mol. Phys. **82**, 383 (1994).

35. M. E. van Leeuwen, Fluid Phase Equilibrium **99**, 1 (1994).

Appendix A

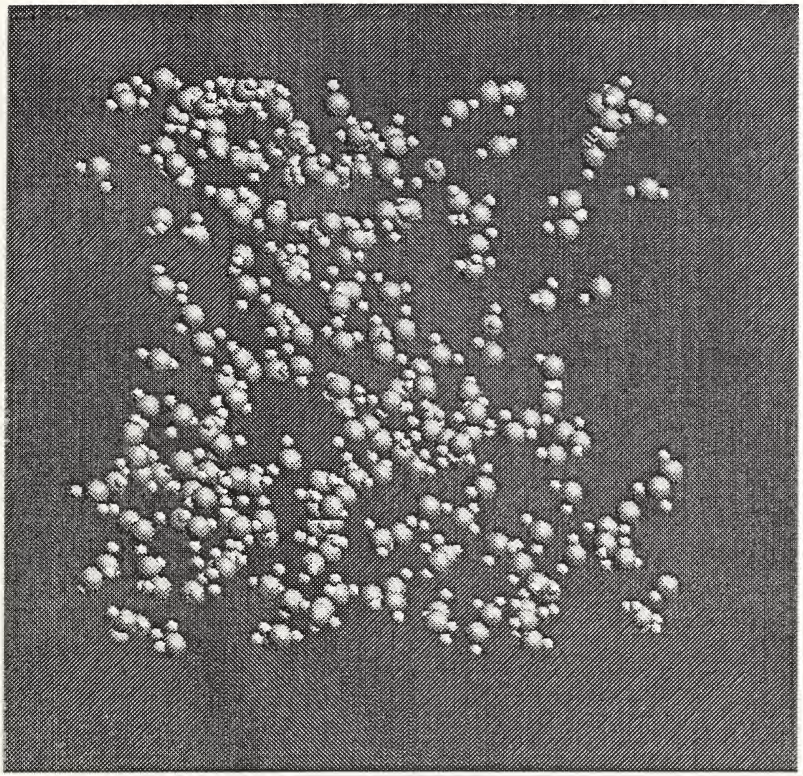
This subroutine determines the number of molecules in each cluster in a configuration. The array `mbtab` contains the labels of the molecules that are bonded to hydrogen sites. Specifically, if hydrogen site *i* on molecule *j* is bonded to the oxygen site on molecule *k*, then `mbtab(j,i) = k`. If that hydrogen site is not bonded to any oxygen site, then `mbtab(j,i) = 0`. In the subroutine, `jcol` refers to a cluster and `krow` refers to a specific member of the cluster. The array `nj` contains the number of molecules in each of the clusters.

```
      subroutine cluster(mbtab,nj)
c
c     A scheme for determining the members of clusters in a set of
c     N particles.
c
      parameter(np=216)
      dimension mbtab(np,2),nj(np)
c       Row j of listc(n,n) is the cluster list for the jth cluster
c       nj(j) is the number of particles in cluster #j.
c       mq(j) is the number of the cluster containing particle j.
      dimension listc(np,np),mq(np)
      do 5 j=1,np
         mq(j)=0
         nj(j)=0
        do 4 k=1,np
            listc(k,j)=0
4       continue
5     continue
c     Construct cluster list, listc.
c     Find the starting point for the first cluster.
      do 910 j=1,np
         do 905 k=1,2
            if(mbtab(j,k).ne.0) then
              go to 920
            end if
905      continue
```

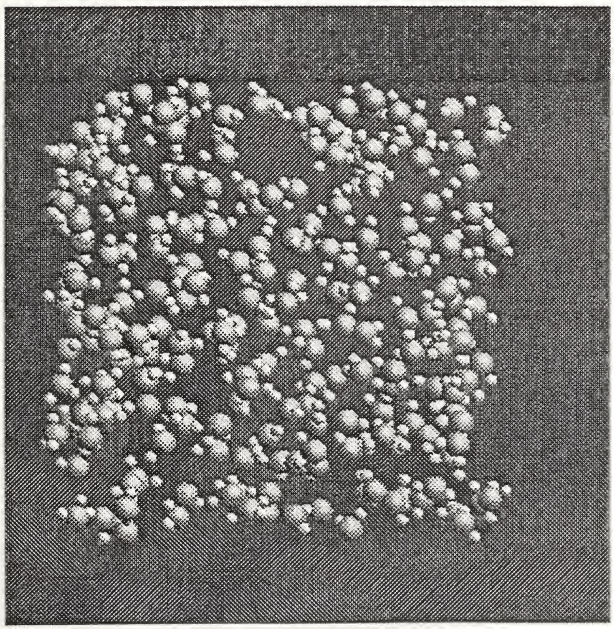
```
910   continue
       STOP 'no clusters!'
920   continue
c
        mt=0
        lrow=1
        jcol=1
        krow=1
200     listc(krow,jcol)=j
          mq(j)=jcol
          mt=mt+1
          nj(jcol)=nj(jcol)+1
100     do 10 jx=1,2
          k=mbtab(j,jx)
          if(k.ne.0 .and.  mq(k).eq.0) then
              krow=krow+1
              listc(krow,jcol)=k
              mq(k)=jcol
              mt=mt+1
              nj(jcol)=nj(jcol)+1
           end if
10       continue
        lrow=lrow+1
        if(lrow.le.krow) then
c     Check on the neighbors of the neighbors of j.
          j=listc(lrow,jcol)
          go to 100
        else
c     Check to see if another cluster is needed.
          if(mt.lt.np) then
              jcol=jcol+1
              krow=1
              lrow=1
c           Find a particle not in a cluster
              do 20 j=1,np
                if(mq(j).eq.0) then
                   go to 200
                end if
20              continue
c             If the program reaches this point---something is wrong!
              STOP 'something is wrong!'
          else
            go to 300
c           No unclustered particles left.
          end if
        end if
300   continue
      return
      end
```

Appendix B

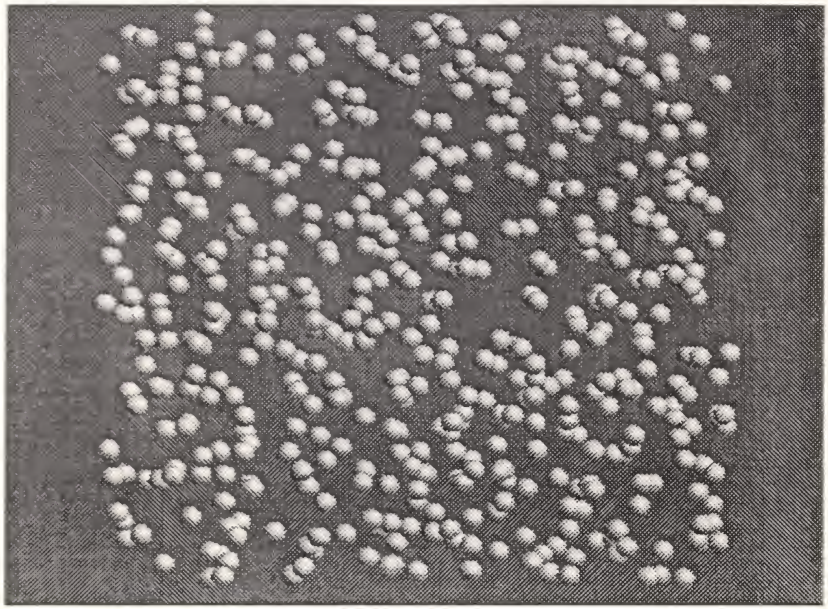
This appendix contains "snapshots" of fluid configurations at temperatures about 10% above the critical temperature. The first one is for SPC/E water at the indicated densities. The second one is for the Lennard-Jones fluid. In both cases, the fluid becomes locally inhomogeneous as the density is decreased.



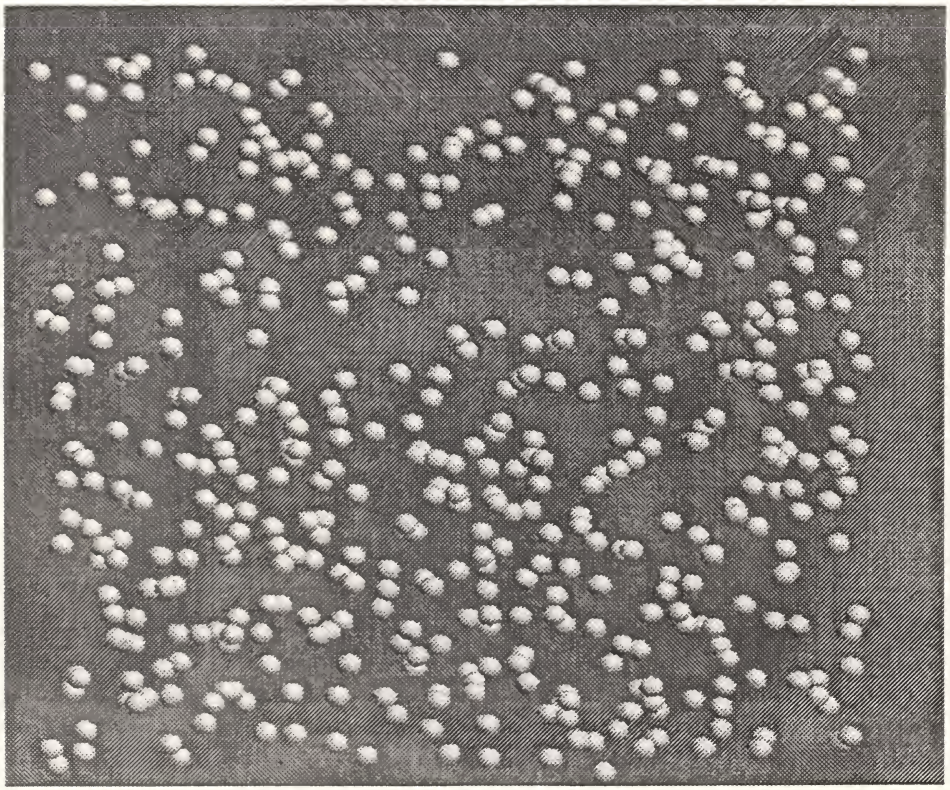
300 kg/m^3



600 kg/m^3



0.4



0.2